

Hutson



1600

RAW SEQUENCE LISTING

DATE: 12/04/2002

PATENT APPLICATION: US/09/642,218

TIME: 14:36:38

Input Set : A:\Pto.amc

Output Set: N:\CRF4\12032002\I642218.raw

SEQUENCE LISTING

ENTERED

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Yurieva, Olga

7 Kuriyan, John

8 O'Donnell, Mike

10 (ii) TITLE OF INVENTION: GAMMA SUBUNIT FOR DNA POLYMERASE

12 (iii) NUMBER OF SEQUENCES: 30

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: David A. Jackson, Esq.

16 (B) STREET: 411 Hackensack Ave, Continental Plaza, 4th

17 Floor

18 (C) CITY: Hackensack

19 (D) STATE: New Jersey

20 (E) COUNTRY: USA

21 (F) ZIP: 07601

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Floppy disk

25 (B) COMPUTER: IBM PC compatible

26 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

27 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

29 (vi) CURRENT APPLICATION DATA:

C--> 30 (A) APPLICATION NUMBER: US/09/642,218

C--> 31 (B) FILING DATE: 18-Aug-2000

32 (C) CLASSIFICATION:

34 (viii) ATTORNEY/AGENT INFORMATION:

35 (A) NAME: Jackson Esq., David A.

36 (B) REGISTRATION NUMBER: 26,742

37 (C) REFERENCE/DOCKET NUMBER: 600-1-179

39 (ix) TELECOMMUNICATION INFORMATION:

40 (A) TELEPHONE: 201-487-5800

41 (B) TELEFAX: 201-343-1684

44 (2) INFORMATION FOR SEQ ID NO: 1:

46 (i) SEQUENCE CHARACTERISTICS:

47 (A) LENGTH: 2007 base pairs

48 (B) TYPE: nucleic acid

49 (C) STRANDEDNESS: double

50 (D) TOPOLOGY: linear

52 (ii) MOLECULE TYPE: DNA (genomic)

54 (iii) HYPOTHETICAL: NO

56 (iv) ANTI-SENSE: NO

61 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

63 TCCGGGGGTG GGGTTCCAG GTAGACCCCG GCCCTCCCG TGAGCCCCTT TACCCAGGCC 60

65 GCCACCTCCT CCAGGGGGGC CAAGGCGTGC AAGGAGAGGA ACGTCCGCAC CACGCCCTAT 120

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/642,218

DATE: 12/04/2002

TIME: 14:36:38

Input Set : A:\Pto.amc

Output Set: N:\CRF4\12032002\I642218.raw

```

67 ACTAGCCTTG TGAGCGCCCT CTACCGCCGC TTCCGCCCCC TCACCTTCCA GGAGGTGGTG 180
69 GGGCAGGAGC ACGTGAAGGA GCCCCTCCTC AAGGCCATCC GGGAGGGGAG GCTCGCCCAG 240
71 GCCTACCTCT TCTCCGGGCC CAGGGGCGTG GGCAAGACCA CCACGGCGAG GCTCCTCGCC 300
73 ATGGCGGTGG GGTGCCAGGG GGAAGACCCC CTTTGGCGGG TCTGCCCCCA CTGCCAGGCG 360
75 GTGCAGAGGG GCGCCCACCC GGACGTGGTG GACATTGACG CCGCCAGCAA CAACTCCGTG 420
77 GAGGACGTGC GGGAGCTGAG GGAAGGATC CACCTCGCCC CCCTCTCTGC CCCCAGGAAG 480
79 GTCTTCATCC TGGACGAGGC CCACATGCTC TCCAAAAGCG CCTTCAACGC CCTCCTCAAG 540
81 ACCCTGGAGG AGCCCCCGCC CCACGTCTCT TTCGTCTTCG CCACCACCGA GCCCAGAGAG 600
83 ATGCCCCCCA CCATCCTCTC CCGCACCCAG CACTTCCGCT TCCGCCGCCT CACGGAGGAG 660
85 GAGATCGCCT TTAAGCTCCG GCGCATCTTG GAGGCCGTGG GGCGGGAGGC GGAGGAGGAG 720
87 GCCCTCCTCC TCCTCGCCCCG CCTGGCGGAC GGGGCCCTTA GGGACGCGGA AAGCCTCCTG 780
89 GAGCGCTTCC TCCTCCTGGA AGGCCCCCTC ACCCGGAAGG AGGTGGAGCG CGCCCTAGGC 840
91 TCCCCCCCAG GGACCGGGGT GGCCGAGATC GCCGCCTCCC TCGCGAGGGG GAAAACGGCG 900
93 GAGGCCCTGG GCCTCGCCCCG GCGCCTCTAC GGGGAAGGGT ACGCCCCGAG GAGCCTGGTC 960
95 TCGGGCCTTT TGGAGGTGTT CCGGGAAGGC CTCTACGCCG CTTTCGGCCT CGCGGGAACC 1020
97 CCCCTTCCCG CCCC GCCCA GGCCTGATC GCCGCCATGA CCGCCCTGGA CGAGGCCATG 1080
99 GAGCGCCTCG CCGCCGCTC CGACGCCTTA AGCCTGGAGG TGGCCCTCCT GGAGGCGGGA 1140
101 AGGGCCCTGG CCGCCGAGGC CCTACCCAG CCCACGGGCG CTCCTTCCCC AGAGGTCCGGC 1200
103 CCCAAGCCGG AAAGCCCCC GACCCCGAA CCCCCAAGGC CCGAGGAGGC GCCCGACCTG 1260
105 CGGGAGCGGT GGCGGGCCTT CCTCGAGGCC CTCAGGCCCA CCCTACGGGC CTTCTGTGCGG 1320
107 GAGGCCCGCC CGGAGGTCCG GGAAGGCCAG CTCTGCCTCG CTTTCCCCGA GGACAAGGCC 1380
109 TTCCACTACC GCAAGGCCTC GGAACAGAAG GTGAGGCTCC TCCCCCTGGC CCAGCCCAT 1440
111 TTCGGGGTGG AGGAGGTCGT CCTCGTCTTG GAGGGAGAAA AAAAAAGCCT GAGCCCAAGG 1500
113 CCCC GCCCGG CCCCACCTCC TGAAGCGCCC GCACCCCGG GCCCTCCCGA GGAGGAGGTA 1560
115 GAGGCGGAGG AAGCGGCGGA GGAGGCCCG GAGGAGGCCT TGAGGCGGGT GGTCCGCCTC 1620
117 CTGGGGGGGC GGGTGCTCTG GGTGCGGCGG CCCAGGACCC GGGAGGCGCC GGAGGAGGAA 1680
119 CCCCTGAGCC AAGACGAGAT AGGGGGTACT GGTATATAAT GGGGGCATGA CGCGGACCAC 1740
121 CGACCTCGGA CAAGAGACCG TGGACAACAT CCTCAAGCGC CTCCGCCGTA TTGAGGGCCA 1800
123 GGTGCGGGGG CTCCAGAAGA TGGTGGCCGA GGGCCGCCCC TCGACGAGG TCCTCACCCA 1860
125 GATGACCGCC ACCAAGAAGG CCATGGAGGC GGCGGCCACC CTGATCCTCC ACGAGTTCTT 1920
127 GAACGTCTGC GCCGCCGAGG TCTCCGAGGG CAAGGTGAAC CCCAAGAAGC CCGAGGAGAT 1980
129 CGCCACCATG CTGAAGAACT TCATCTA 2007

```

131 (2) INFORMATION FOR SEQ ID NO: 2:

133 (i) SEQUENCE CHARACTERISTICS:

134 (A) LENGTH: 529 amino acids

135 (B) TYPE: amino acid

136 (C) STRANDEDNESS: single

137 (D) TOPOLOGY: linear

139 (ii) MOLECULE TYPE: protein

141 (iii) HYPOTHETICAL: NO

143 (v) FRAGMENT TYPE: internal

148 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

150 Met Ser Ala Leu Tyr Arg Arg Phe Arg Pro Leu Thr Phe Gln Glu Val

151 1 5 10 15

153 Val Gly Gln Glu His Val Lys Glu Pro Leu Leu Lys Ala Ile Arg Glu

154 20 25 30

156 Gly Arg Leu Ala Gln Ala Tyr Leu Phe Ser Gly Pro Arg Gly Val Gly

157 35 40 45

159 Lys Thr Thr Thr Ala Arg Leu Leu Ala Met Ala Val Gly Cys Gln Gly

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/642,218

DATE: 12/04/2002

TIME: 14:36:38

Input Set : A:\Pto.amc

Output Set: N:\CRF4\12032002\I642218.raw

160	50	55	60
162	Glu Asp Pro Pro Cys Gly Val Cys Pro His Cys Gln Ala Val Gln Arg		
163	65	70	75
165	Gly Ala His Pro Asp Val Val Asp Ile Asp Ala Ala Ser Asn Asn Ser		80
166		85	90
168	Val Glu Asp Val Arg Glu Leu Arg Glu Arg Ile His Leu Ala Pro Leu		95
169		100	105
171	Ser Ala Pro Arg Lys Val Phe Ile Leu Asp Glu Ala His Met Leu Ser		110
172		115	120
174	Lys Ser Ala Phe Asn Ala Leu Leu Lys Thr Leu Glu Glu Pro Pro Pro		125
175		130	135
177	His Val Leu Phe Val Phe Ala Thr Thr Glu Pro Glu Arg Met Pro Pro		140
178		145	150
180	Thr Ile Leu Ser Arg Thr Gln His Phe Arg Phe Arg Arg Leu Thr Glu		155
181		165	170
183	Glu Glu Ile Ala Phe Lys Leu Arg Arg Ile Leu Glu Ala Val Gly Arg		175
184		180	185
186	Glu Ala Glu Glu Glu Ala Leu Leu Leu Leu Ala Arg Leu Ala Asp Gly		190
187		195	200
189	Ala Leu Arg Asp Ala Glu Ser Leu Leu Glu Arg Phe Leu Leu Leu Glu		205
190		210	215
192	Gly Pro Leu Thr Arg Lys Glu Val Glu Arg Ala Leu Gly Ser Pro Pro		220
193		225	230
195	Gly Thr Gly Val Ala Glu Ile Ala Ala Ser Leu Ala Arg Gly Lys Thr		235
196		245	250
198	Ala Glu Ala Leu Gly Leu Ala Arg Arg Leu Tyr Gly Glu Gly Tyr Ala		255
199		260	265
201	Pro Arg Ser Leu Val Ser Gly Leu Leu Glu Val Phe Arg Glu Gly Leu		270
202		275	280
204	Tyr Ala Ala Phe Gly Leu Ala Gly Thr Pro Leu Pro Ala Pro Pro Gln		285
205		290	295
207	Ala Leu Ile Ala Ala Met Thr Ala Leu Asp Glu Ala Met Glu Arg Leu		300
208		305	310
210	Ala Arg Arg Ser Asp Ala Leu Ser Leu Glu Val Ala Leu Leu Glu Ala		315
211		325	330
213	Gly Arg Ala Leu Ala Ala Glu Ala Leu Pro Gln Pro Thr Gly Ala Pro		335
214		340	345
216	Ser Pro Glu Val Gly Pro Lys Pro Glu Ser Pro Pro Thr Pro Glu Pro		350
217		355	360
219	Pro Arg Pro Glu Glu Ala Pro Asp Leu Arg Glu Arg Trp Arg Ala Phe		365
220		370	375
222	Leu Glu Ala Leu Arg Pro Thr Leu Arg Ala Phe Val Arg Glu Ala Arg		380
223		385	390
225	Pro Glu Val Arg Glu Gly Gln Leu Cys Leu Ala Phe Pro Glu Asp Lys		395
226		405	410
228	Ala Phe His Tyr Arg Lys Ala Ser Glu Gln Lys Val Arg Leu Leu Pro		415
229		420	425
231	Leu Ala Gln Ala His Phe Gly Val Glu Glu Val Val Leu Val Leu Glu		430
232		435	440
			445

RAW SEQUENCE LISTING

DATE: 12/04/2002

PATENT APPLICATION: US/09/642,218

TIME: 14:36:38

Input Set : A:\Pto.amc

Output Set: N:\CRF4\12032002\I642218.raw

```

234   Gly Glu Lys Lys Ser Leu Ser Pro Arg Pro Arg Pro Ala Pro Pro Pro
235           450                     455                     460
237   Glu Ala Pro Ala Pro Pro Gly Pro Pro Glu Glu Val Glu Ala Glu
238   465                     470                     475                     480
240   Glu Ala Ala Glu Glu Ala Pro Glu Glu Ala Leu Arg Arg Val Val Arg
241           485                     490                     495
243   Leu Leu Gly Gly Arg Val Leu Trp Val Arg Arg Pro Arg Thr Arg Glu
244           500                     505                     510
246   Ala Pro Glu Glu Glu Pro Leu Ser Gln Asp Glu Ile Gly Gly Thr Gly
247           515                     520                     525
249   Ile

```

252 (2) INFORMATION FOR SEQ ID NO: 3:

254 (i) SEQUENCE CHARACTERISTICS:

255 (A) LENGTH: 1590 base pairs

256 (B) TYPE: nucleic acid

257 (C) STRANDEDNESS: double

258 (D) TOPOLOGY: linear

260 (ii) MOLECULE TYPE: DNA (genomic)

262 (iii) HYPOTHETICAL: NO

264 (iv) ANTI-SENSE: NO

269 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```

271 GTGAGCGCCC TCTACGCGCG CTTCCGCCCC CTCACCTTCC AGGAGGTGGT GGGGCAGGAG      60
273 CACGTGAAGG AGCCCCTCCT CAAGGCCATC CGGGAGGGGA GGCTCGCCCA GGCCTACCTC      120
275 TTCTCCGGGC CCAGGGGCGT GGGCAAGACC ACCACGGCGA GGCTCCTCGC CATGGCGGTG      180
277 GGGTGCCAGG GGAAGACCC CCCTTGCGGG GTCTGCCCCC ACTGCCAGGC GGTGCAGAGG      240
279 GGCGCCCACC CGGACGTGGT GGACATTGAC GCCGCCAGCA ACAACTCCGT GGAGGACGTG      300
281 CGGGAGCTGA GGAAGAGGAT CCACCTCGCC CCCCTCTCTG CCCCAGGAA GGTCTTCATC      360
283 CTGGACGAGG CCCACATGCT CTCCAAAAGC GCCTTCAACG CCCTCCTCAA GACCCTGGAG      420
285 GAGCCCCCGC CCCACGTCCT CTTCTGTTT GCCACCACCG AGCCCGAGAG GATGCCCCCC      480
287 ACCATCCTCT CCCGCACCCA GCACTTCCGC TTCCGCGGCC TCACGGAGGA GGAGATCGCC      540
289 TTTAAGCTCC GGCGCATCCT GGAGGCCGTG GGGCGGGAGG CGGAGGAGGA GGCCCTCCTC      600
291 CTCCTCGCCC GCCTGGCGGA CGGGGCCCTT AGGGACGCGG AAAGCCTCCT GGAGCGCTTC      660
293 CTCCTCCTGG AAGGCCCCCT CACCCGGAAG GAGGTGGAGC GCGCCCTAGG CTCCCCCCA      720
295 GGGACCGGGG TGGCCGAGAT CGCCGCCTCC CTCGCGAGGG GGAAAACGGC GGAGGCCCTG      780
297 GGCTCGCCC GGCGCCTCTA CGGGGAAGGG TACGCCCCGA GGAGCCTGGT CTCGGGCCTT      840
299 TTGGAGGTGT TCCGGGAAGG CCTCTACGCC GCCTTCGGCC TCGCGGGAAC CCCCCTTCCC      900
301 GCCCCGCCCC AGGCCCTGAT CGCCGCCATG ACCGCCCTGG ACGAGGCCAT GGAGCGCCTC      960
303 GCCCCGCCGT CCGACGCCTT AAGCCTGGAG GTGGCCCTCC TGGAGGCGGG AAGGGCCCTG     1020
305 GCCGCCGAGG CCCTACCCCA GCCCACGGGC GCTCCTTCCC CAGAGGTCGG CCCCAGCCG     1080
307 GAAAGCCCCC CGACCCCGGA ACCCCCAAGG CCCGAGGAGG CGCCCGACCT GCGGGAGCGG     1140
309 TGGCGGGCCT TCCTCGAGGC CCTCAGGCC ACCCTACGGG CCTTCGTGCG GGAGGCCCGC     1200
311 CCGGAGGTCC GGAAGGCCA GCTCTGCCTC GCTTTCCCCG AGGACAAGGC CTTCCACTAC     1260
313 CGCAAGGCCT CGGAACAGAA GGTGAGGCTC CTCCCCCTGG CCCAGGCCA TTTCGGGGTG     1320
315 GAGGAGGTCT TCCTCGTCCT GGAGGGAGAA AAAAAAAGCC TGAGCCCAAG GCCCCGCCG     1380
317 GCCCCACCTC CTGAAGCGCC CGCACCCCGG GGCCCTCCCG AGGAGGAGGT AGAGGCGGAG     1440
319 GAAGCGGCGG AGGAGGCCCC GGAGGAGGCC TTGAGGCGGG TGGTCCGCTT CCTGGGGGGG     1500
321 CGGGTGCTCT GGGTGCGGCG GCCCAGGACC CGGGAGGCGC CGGAGGAGGA ACCCCTGAGC     1560
323 CAAGACGAGA TAGGGGGTAC TGGTATATAA
325 (2) INFORMATION FOR SEQ ID NO: 4:

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/642,218

DATE: 12/04/2002

TIME: 14:36:38

Input Set : A:\Pto.amc

Output Set: N:\CRF4\12032002\I642218.raw

```

327 (i) SEQUENCE CHARACTERISTICS:
328 (A) LENGTH: 464 amino acids
329 (B) TYPE: amino acid
330 (C) STRANDEDNESS: single
331 (D) TOPOLOGY: linear
333 (ii) MOLECULE TYPE: protein
335 (iii) HYPOTHETICAL: NO
337 (v) FRAGMENT TYPE: internal
342 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
344 Met Ser Ala Leu Tyr Arg Arg Phe Arg Pro Leu Thr Phe Gln Glu Val
345 1 5 10 15
347 Val Gly Gln Glu His Val Lys Glu Pro Leu Leu Lys Ala Ile Arg Glu
348 20 25 30
350 Gly Arg Leu Ala Gln Ala Tyr Leu Phe Ser Gly Pro Arg Gly Val Gly
351 35 40 45
353 Lys Thr Thr Thr Ala Arg Leu Leu Ala Met Ala Val Gly Cys Gln Gly
354 50 55 60
356 Glu Asp Pro Pro Cys Gly Val Cys Pro His Cys Gln Ala Val Gln Arg
357 65 70 75 80
359 Gly Ala His Pro Asp Val Val Asp Ile Asp Ala Ala Ser Asn Asn Ser
360 85 90 95
362 Val Glu Asp Val Arg Glu Leu Arg Glu Arg Ile His Leu Ala Pro Leu
363 100 105 110
365 Ser Ala Pro Arg Lys Val Phe Ile Leu Asp Glu Ala His Met Leu Ser
366 115 120 125
368 Lys Ser Ala Phe Asn Ala Leu Leu Lys Thr Leu Glu Glu Pro Pro Pro
369 130 135 140
371 His Val Leu Phe Val Phe Ala Thr Thr Glu Pro Glu Arg Met Pro Pro
372 145 150 155 160
374 Thr Ile Leu Ser Arg Thr Gln His Phe Arg Phe Arg Arg Leu Thr Glu
375 165 170 175
377 Glu Glu Ile Ala Phe Lys Leu Arg Arg Ile Leu Glu Ala Val Gly Arg
378 180 185 190
380 Glu Ala Glu Glu Glu Ala Leu Leu Leu Ala Arg Leu Ala Asp Gly
381 195 200 205
383 Ala Leu Arg Asp Ala Glu Ser Leu Leu Glu Arg Phe Leu Leu Leu Glu
384 210 215 220
386 Gly Pro Leu Thr Arg Lys Glu Val Glu Arg Ala Leu Gly Ser Pro Pro
387 225 230 235 240
389 Gly Thr Gly Val Ala Glu Ile Ala Ala Ser Leu Ala Arg Gly Lys Thr
390 245 250 255
392 Ala Glu Ala Leu Gly Leu Ala Arg Arg Leu Tyr Gly Glu Gly Tyr Ala
393 260 265 270
395 Pro Arg Ser Leu Val Ser Gly Leu Leu Glu Val Phe Arg Glu Gly Leu
396 275 280 285
398 Tyr Ala Ala Phe Gly Leu Ala Gly Thr Pro Leu Pro Ala Pro Pro Gln
399 290 295 300
401 Ala Leu Ile Ala Ala Met Thr Ala Leu Asp Glu Ala Met Glu Arg Leu
402 305 310 315 320

```